

Inferring protein interaction from sequence co-evolution

Sequence variations, protein interactions, and diseases



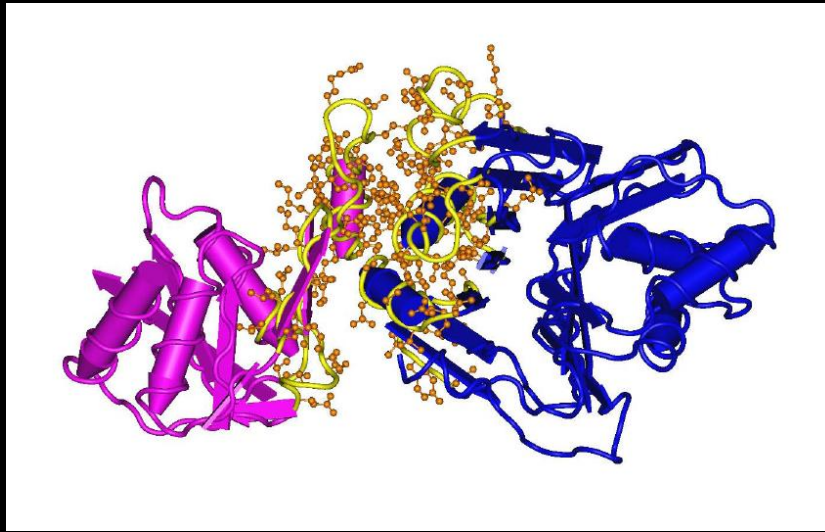
Teresa Przytycka
NIH / NLM / NCBI



October 7, 2009

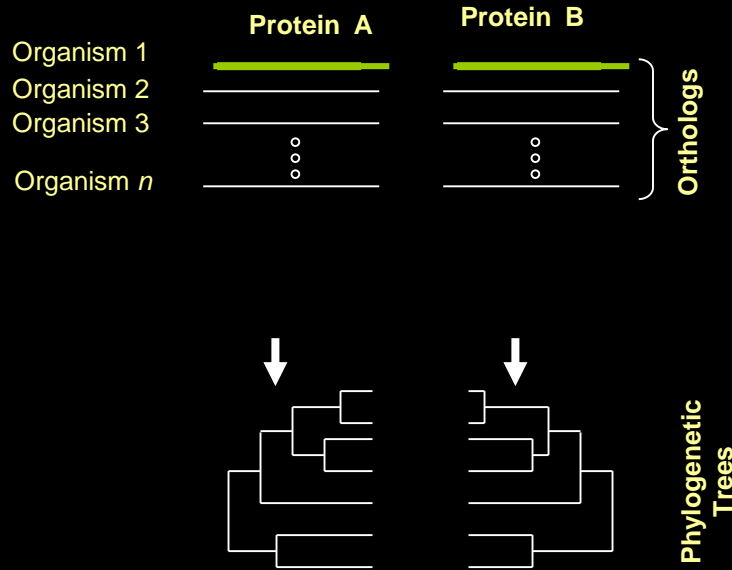
- Protein interaction and sequence co-evolution

Interacting proteins are expected to co-evolve to ensure proper binding

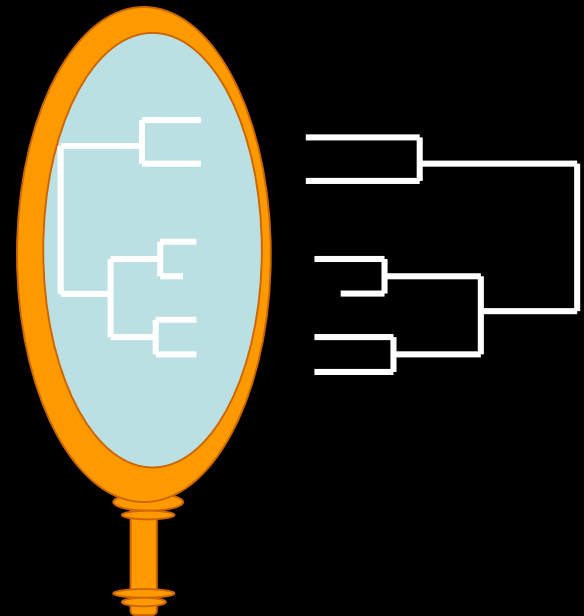


Inferring protein interaction from the co-evolution principle

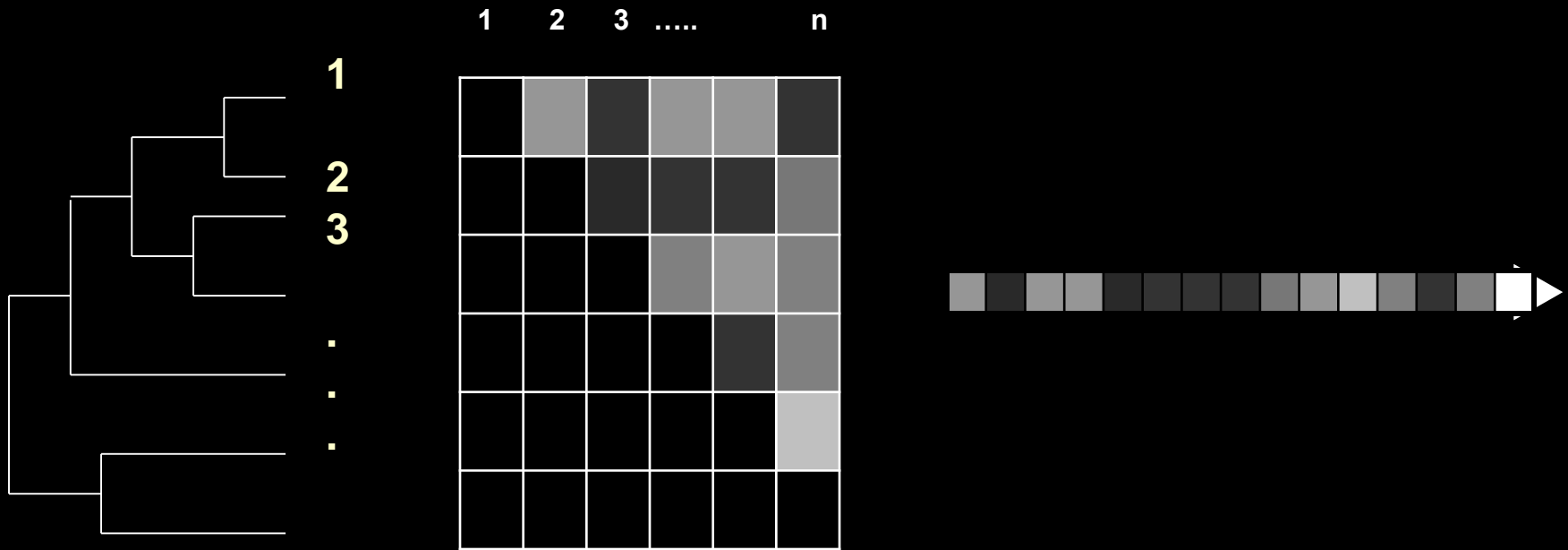
[Goh et al 2000, Pazos and Valencia 2001]



Mirrortree Method:

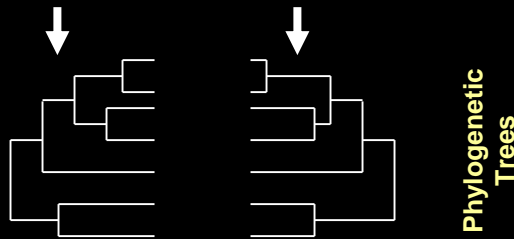
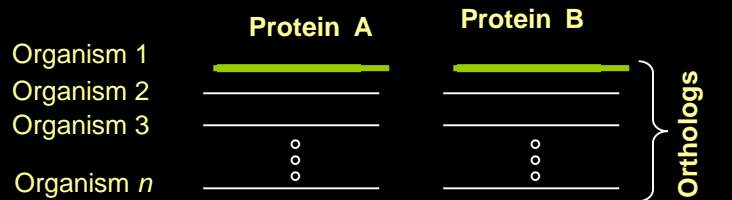


Evolutionary vector

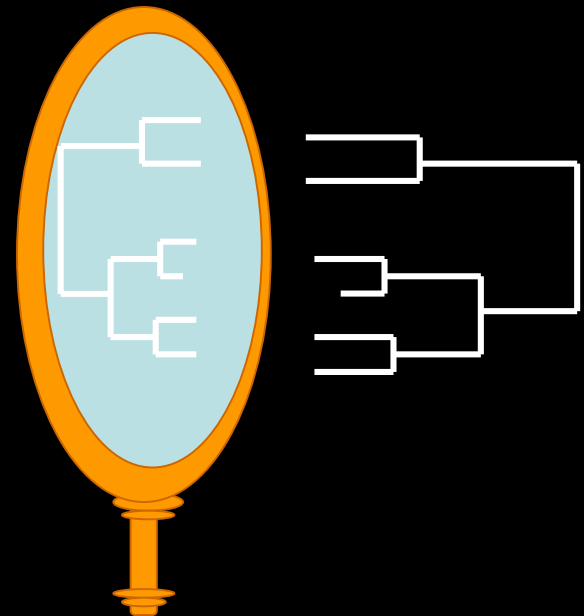


Distance Matrix

Inferring protein interaction from the co-evolution principle



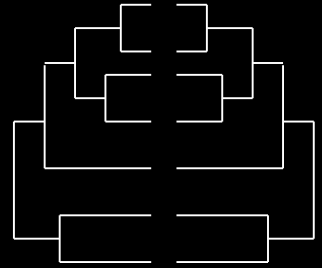
Mirrortree Method:



Compute correlation

Simple idea but lot's of questions...

- How to separate co-evolution due to common speciation history from co-evolution due to function?



Kann et. al. *Proteins* 2007

- Is the co- evolution signal distributed uniformly over the sequence? Between binding site only?

Kann et. al. *JMB* 2009

Jothi et. al. *JMB* 2007

- Predicting Interaction specificity

Jothi et. al. *Bioinformatics* 2005

Challenges

- How to separate co-evolution due to common speciation history from co-evolution due to function?

Kann et. al. *Proteins* 2007

- Is the co- evolution signal distributed uniformly over the sequence? Between binding site only?

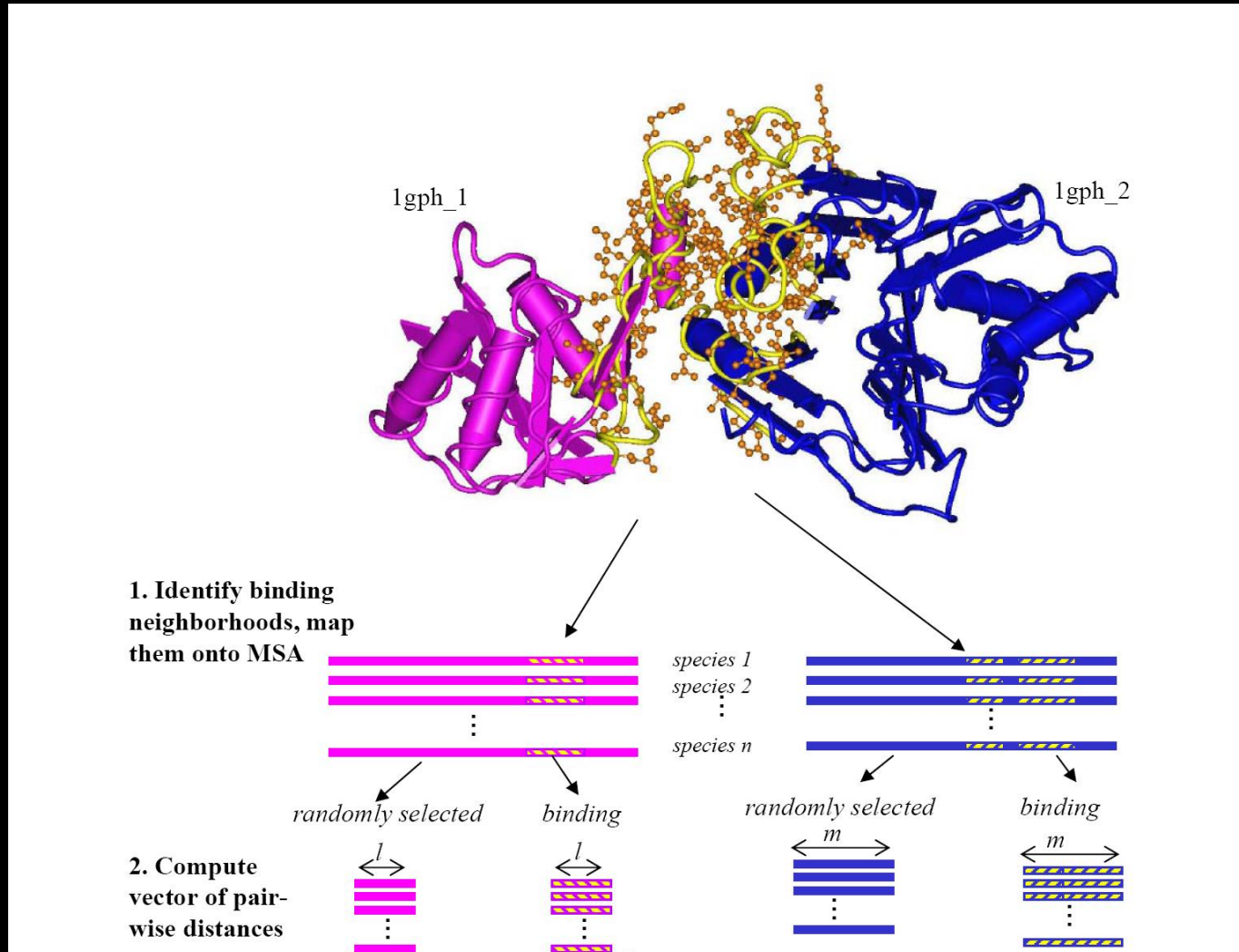
Kann et. al. *JMB* 2009

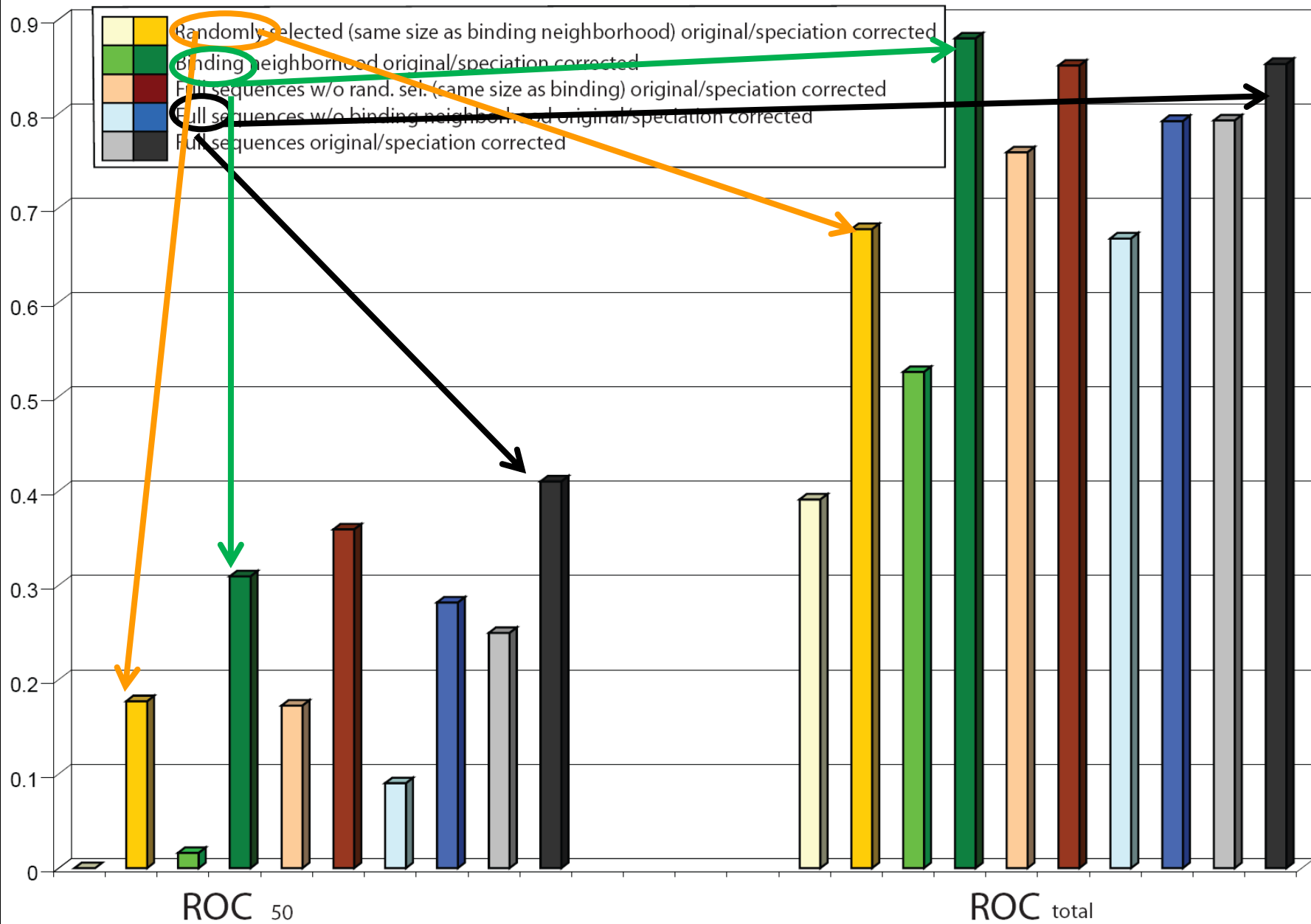
Jothi et. al. *JMB* 2007

- Predicting Interaction specificity

Jothi et. al. *Bioinformatics* 2005

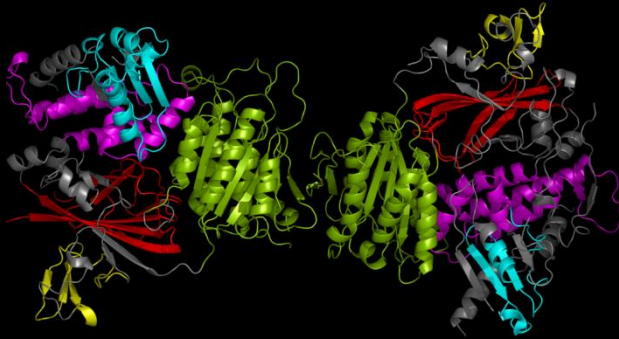
Do binding sites co-evolve more tightly?





Binding sites are important but not the only contributor of the co- evolutionary signal

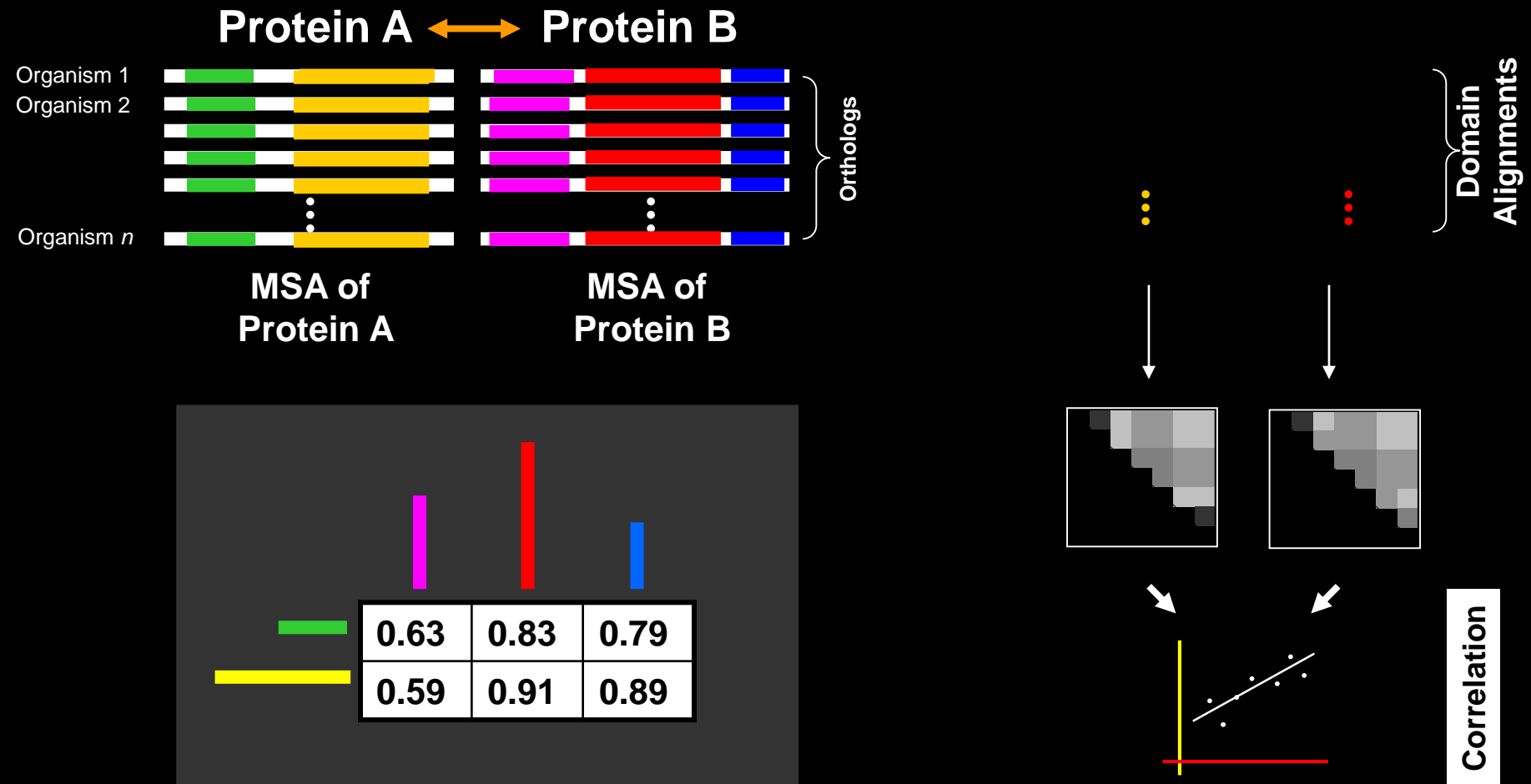
Predicting interacting domains



**Given interacting multi-domain proteins
domains that are in contact**

Jothi et. al. *JMB* 2007

Mirror tree approach can be used to recognize interacting domains

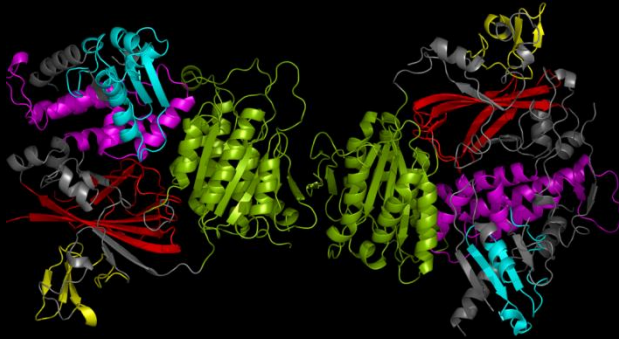


RESULTS:

In 64% cases, the domain pair with highest correlation was interacting (55% expected by chance)

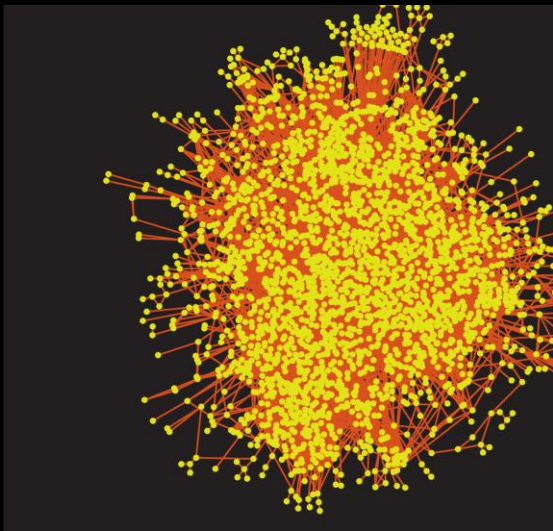
Jothi et. al. *JMB* 2007

Predicting interacting domains



Given interacting multi-domain proteins
domains that are in contact

Jothi et. al. *JMB* 2007

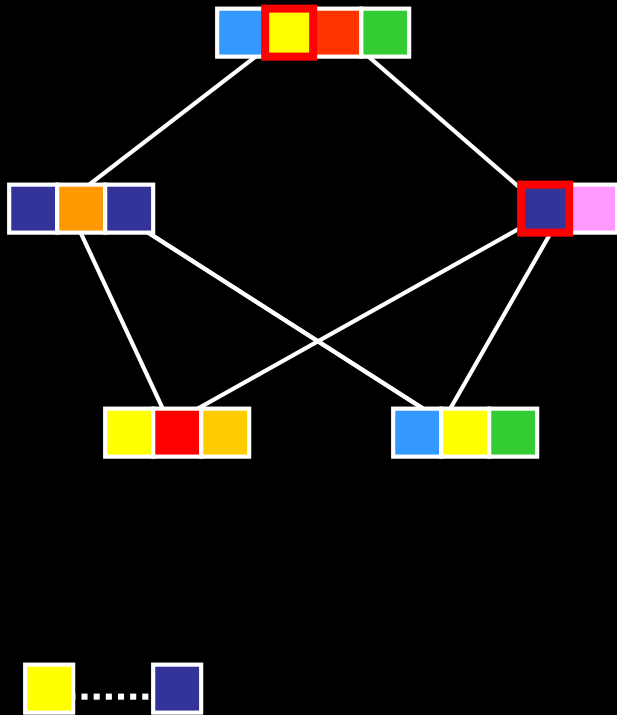


Given protein-protein interaction network
domains that are in contact

Guimaraes et. al. *Genome Biology* 2008

- Protein interaction and sequence co-evolution
- Predicting domain interaction from protein interaction networks

Parsimony approach

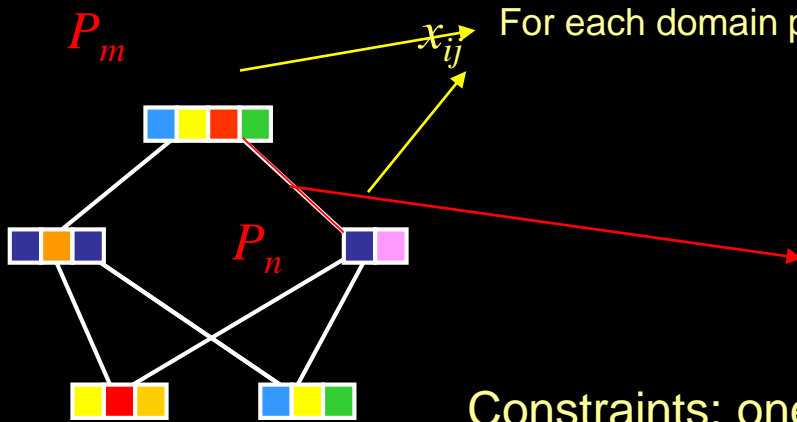


Assumption: Protein interactions are mediated by domain interactions

Hypothesis: Interactions evolved in most parsimonious way

Method: Find the smallest set of domain pairs whose interaction would explain all protein interactions in the network

Linear programming formulation



For each domain pair D_i, D_j : variable x_{ij} taking value 0 or 1

$$\sum_{(D_i, D_j) \in (P_m, P_n)} x_{ij} \geq 1$$

Constraints: one per protein interaction P_m, P_n

Objective function
(representing parsimony assumption):

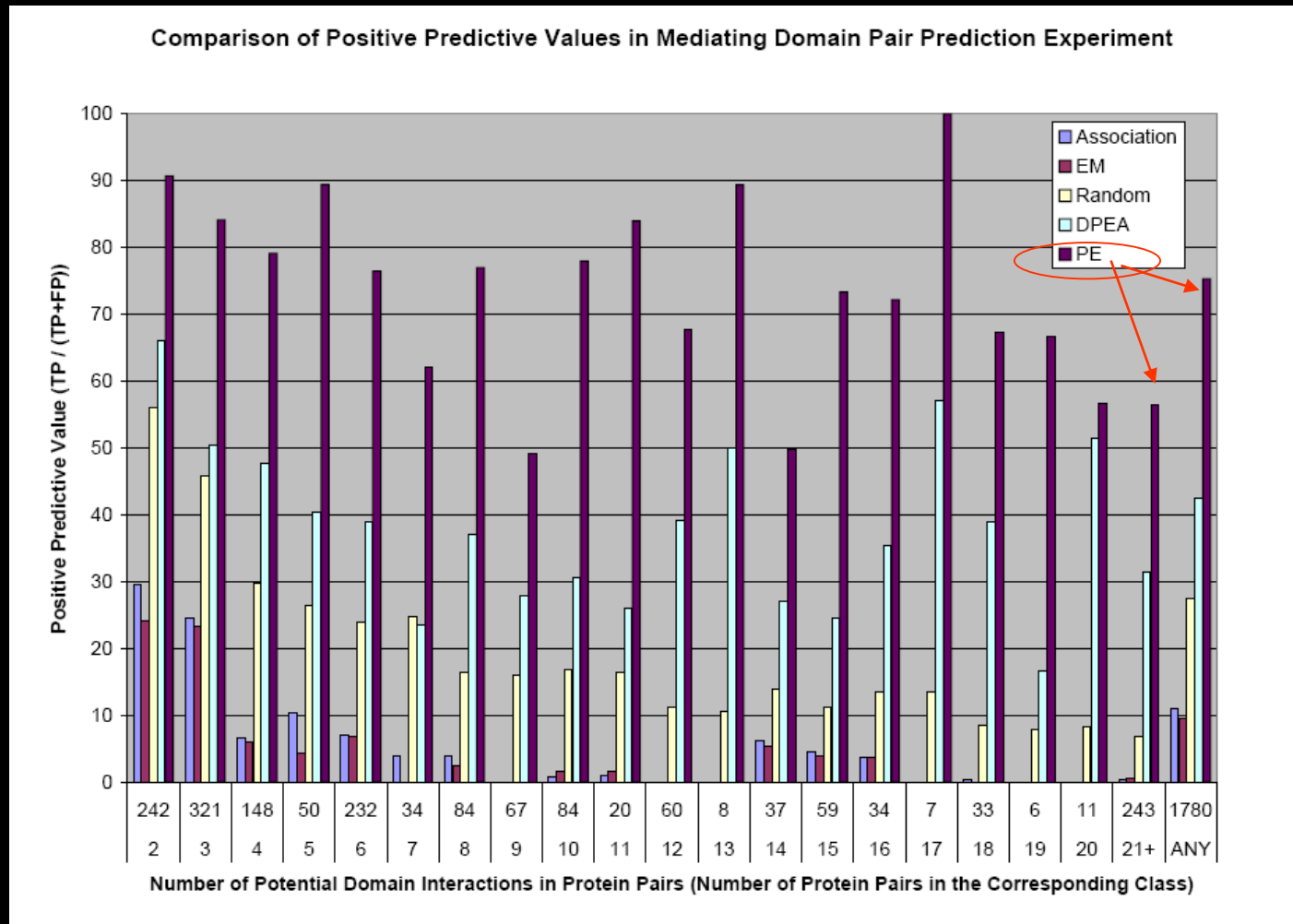
$$\text{minimize} \quad \sum_{D_i, D_j} x_{ij}$$

Interacting domains pairs – domains pairs with $x_{ij}=1$

Additional problems to solve:

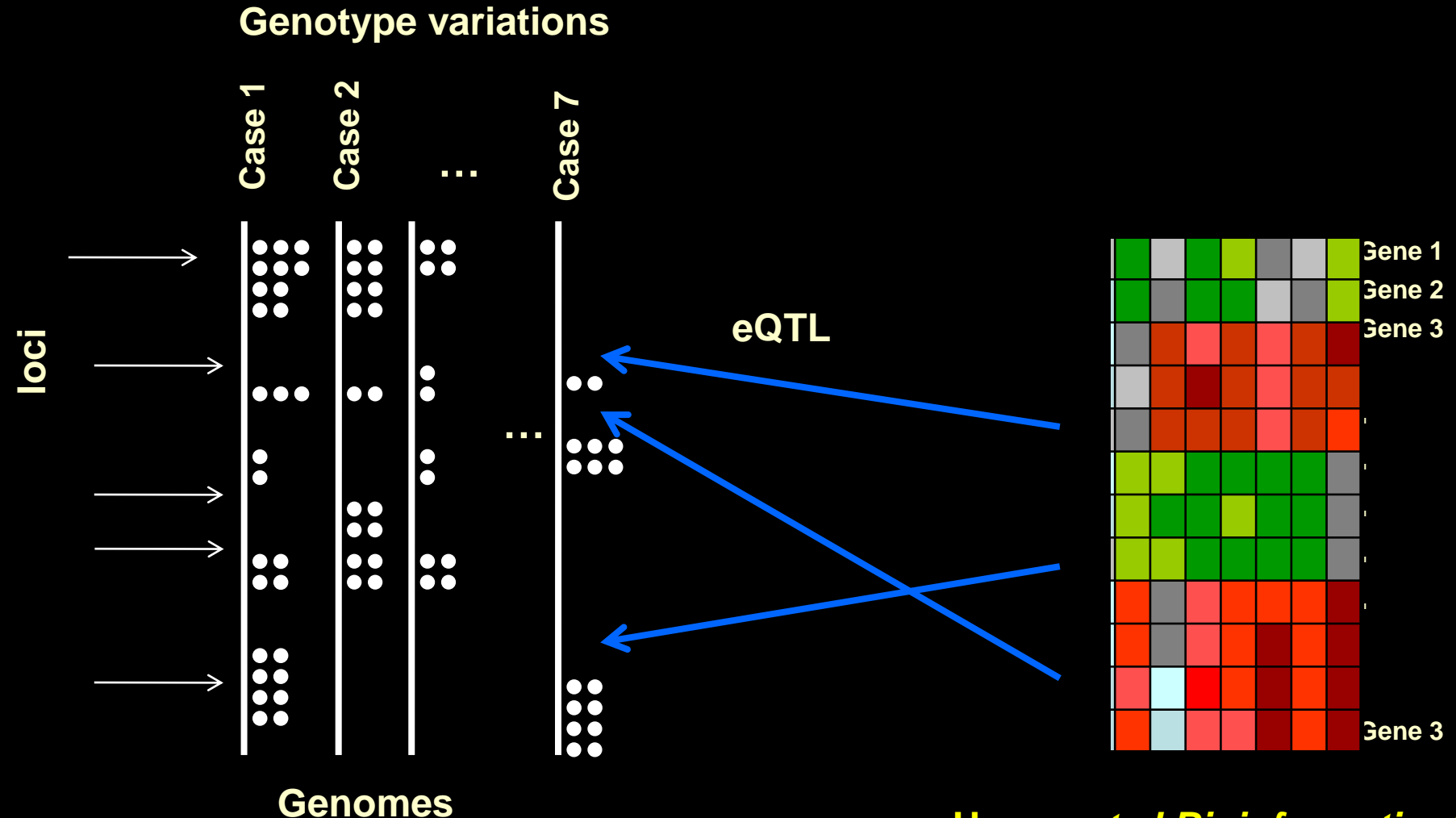
- Model the noise in the network
- Estimate p-values

Results compared to previous methods: Identifying interacting domain pair in interacting protein pair

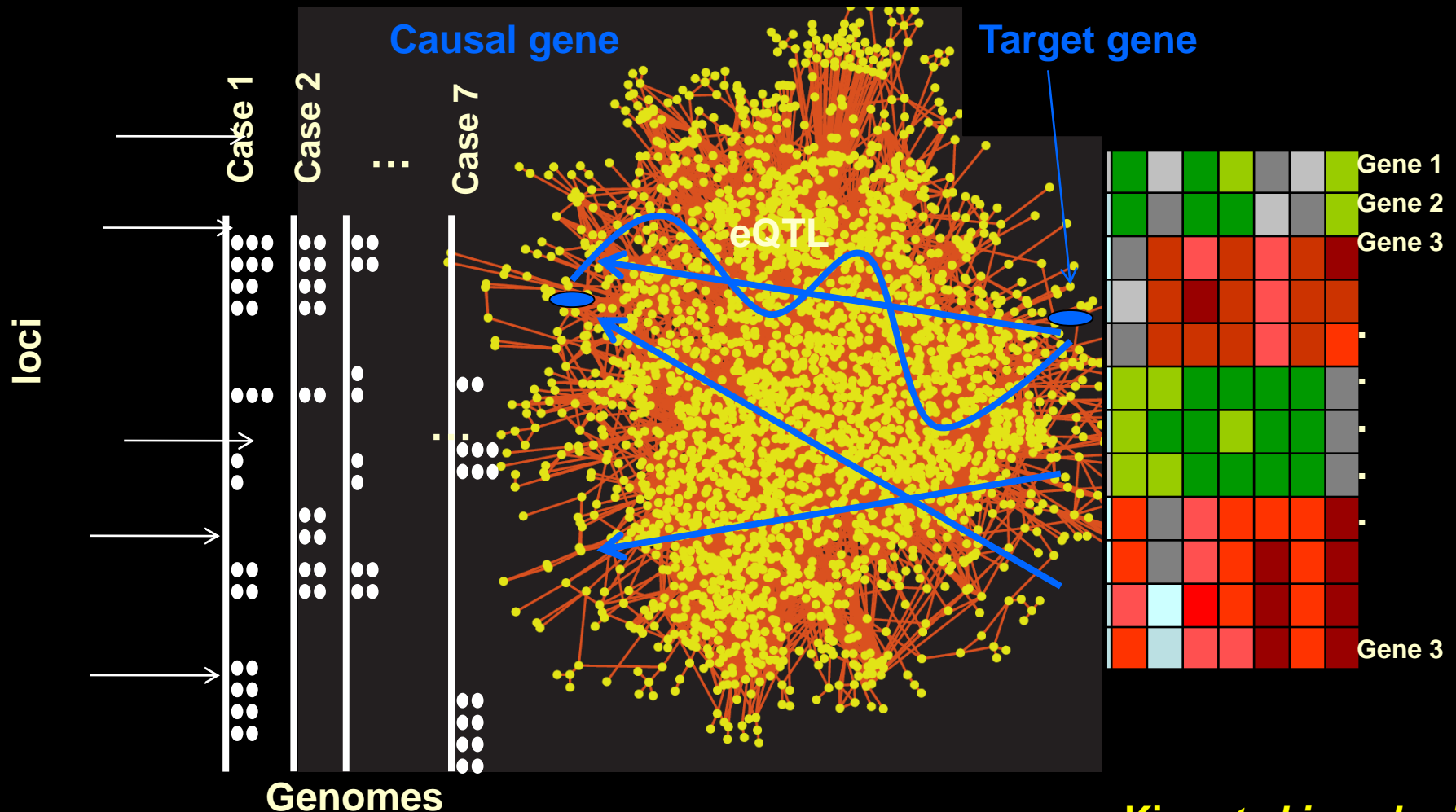


- Protein interaction and sequence co-evolution
- Predicting domain interaction from protein interaction networks
- Combining genetic sequence variation, genome wide expression profile and protein interaction to infer pathways dysregulated in complex diseases

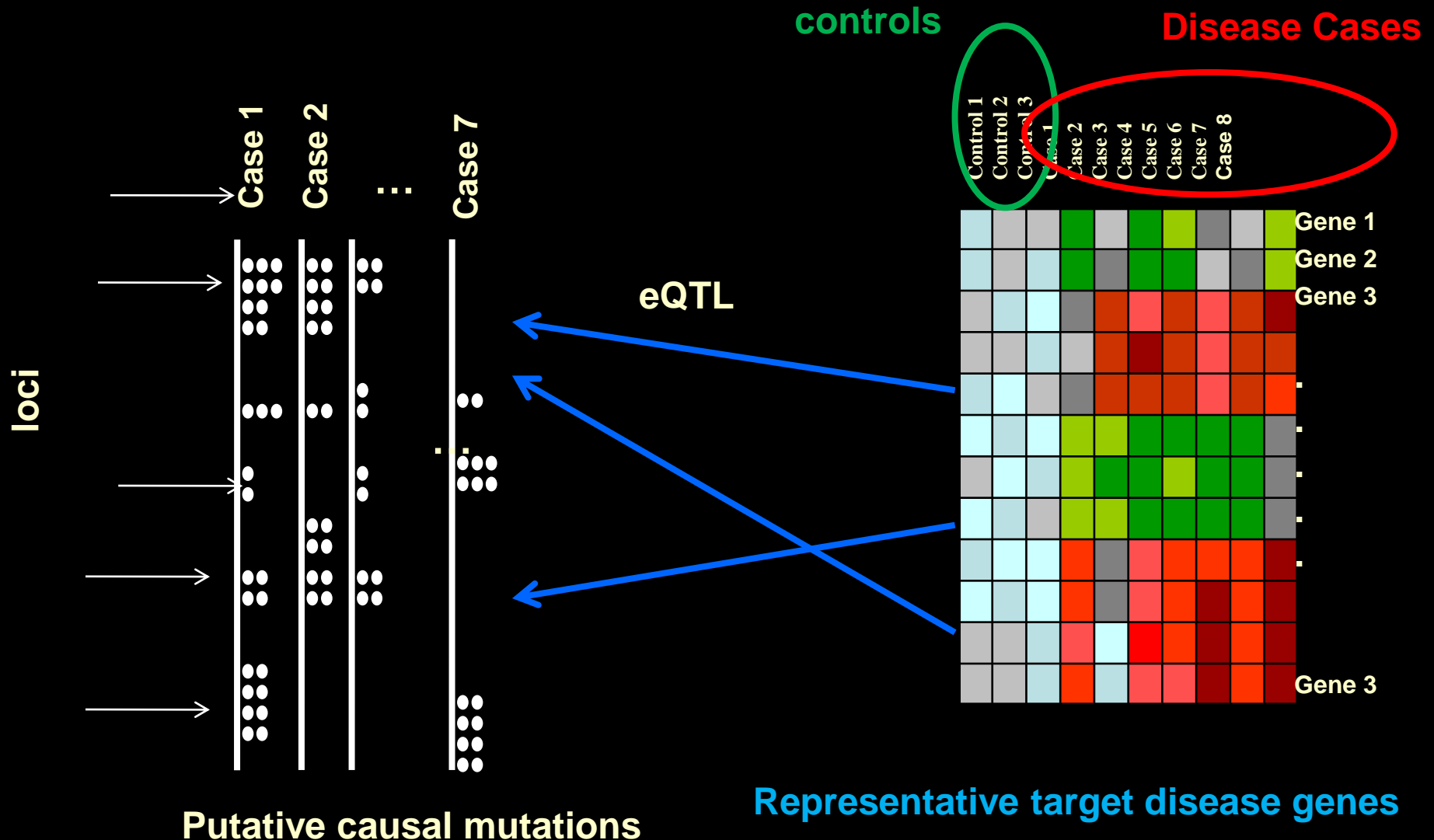
Genetic variations in individuals affects gene expression level



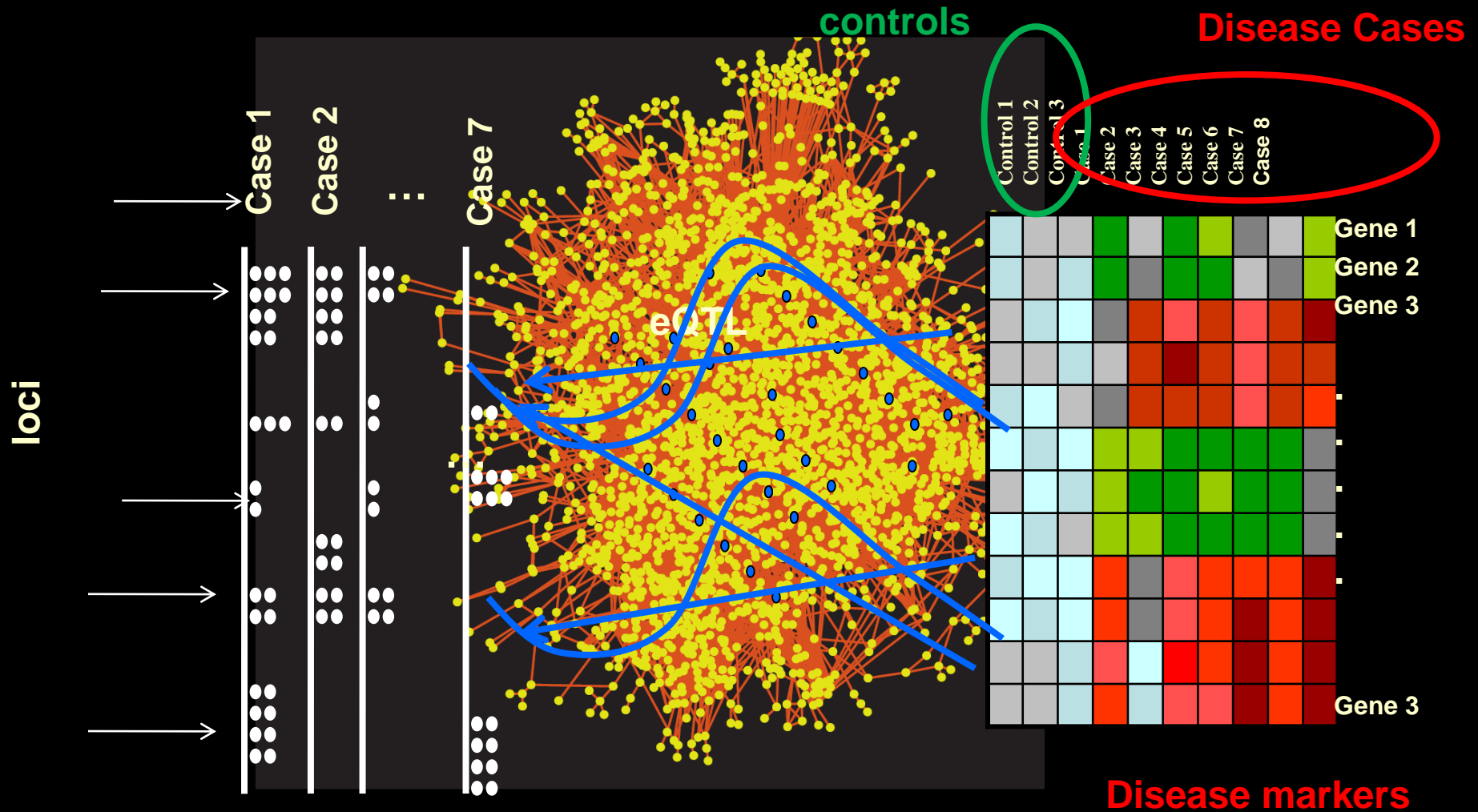
Bringing PPI network and other high throughput networks



Associations of genes expressed differently in disease /control groups are primary target



Uncovering causal genes and dys-regulated pathways



Causal genes

Kim et.al. submitted

Acknowledgments

Former lab members

Katia Guimaraes (associate professor, Brazil)

Raja Jothi (currently PI at NIEHS)

Elena Zotenko (currently Max Planck Institute)

Current lab members

Dong Yeon Cho

Yoo-ah Kim

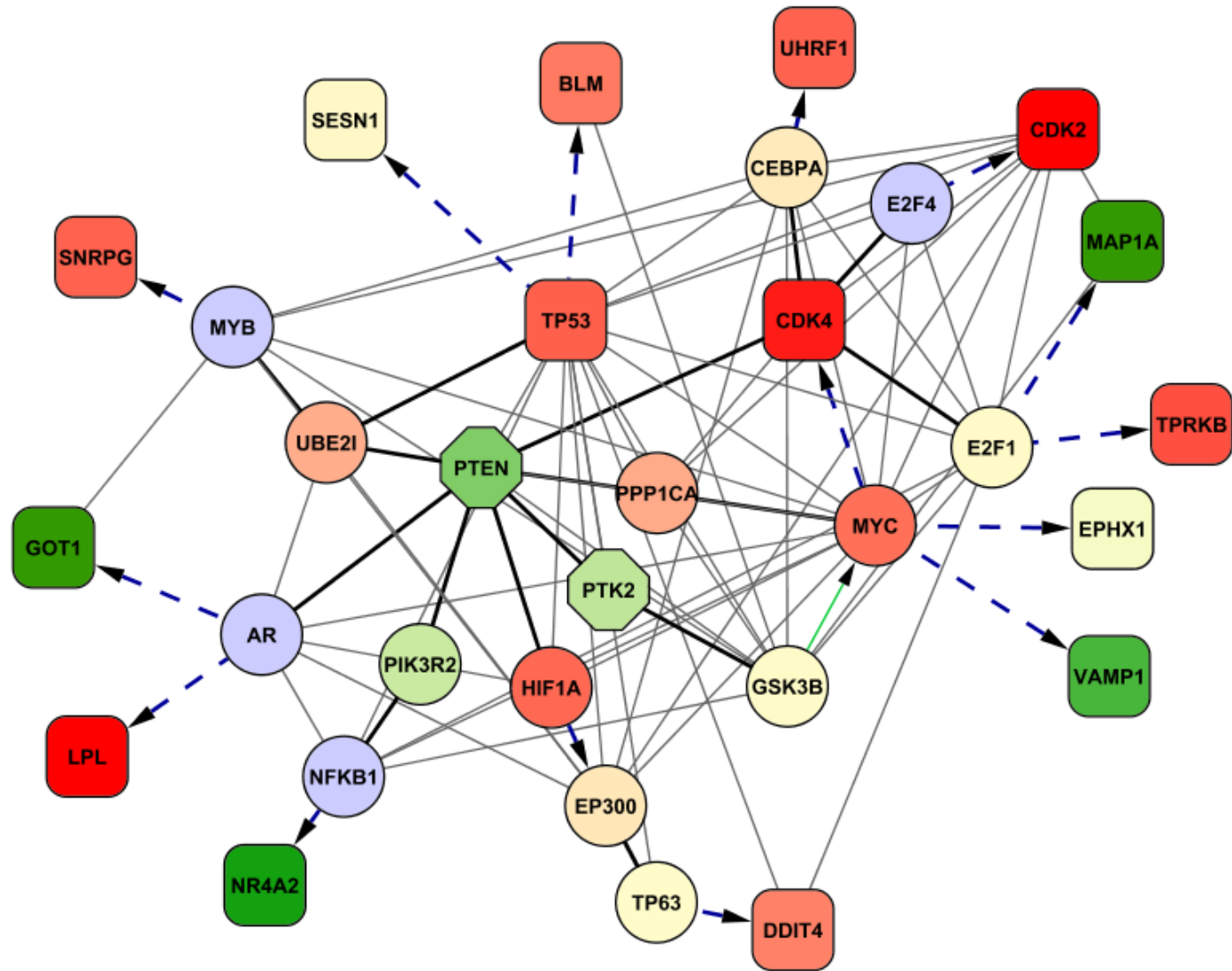
Yang Huang

Damian Wojtowicz

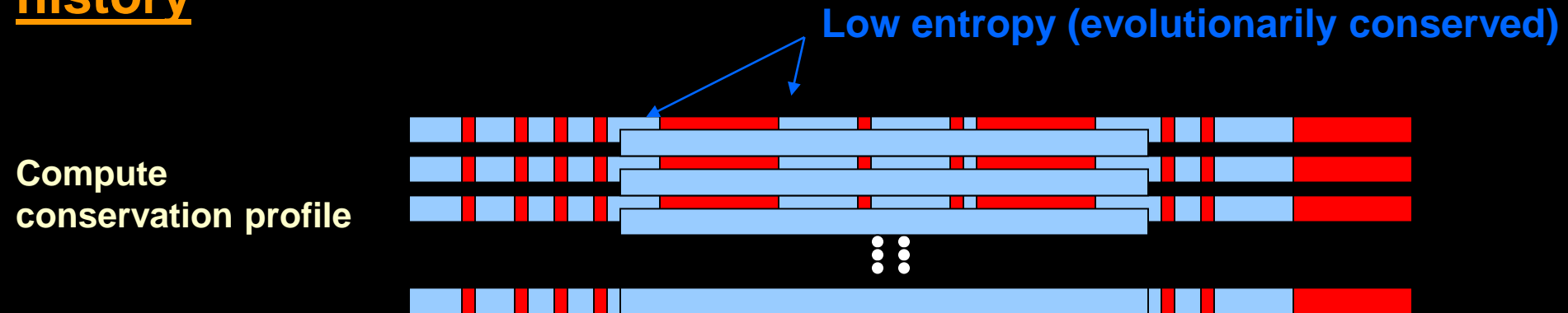
Jie Zheng

Collaborators

Maricel Kann UMBC



Evolutionarily conserved regions help separate functional co-evolution from co-evolution due common speciation history



Use conserved positions only

Additional correction using previously mentioned methods

